

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2000, 08:48:24 : Search time 378.27 Seconds  
(without alignments)  
94.582 Million cell updated/sec

Title: US-09-428-674-11

Perfect score: 143  
Sequence: 1 gtggcctatgtactgtccta.....aattgtttatatttgcgtt 143

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size: 15

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N.Geneseq.36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	14.7	84	1 T21573	Human gene signatu
2	17	11.9	4121	1 V74626	Staphylococcus aur
3	17	11.9	5790	1 X2464	Bovine myostatin g
4	17	11.9	6029	1 X23824	Bacillus sp. GUS 9
5	16	11.2	200	1 Q77245	Human genome fragm
6	16	11.2	211	1 T19250	Human gene signatu
7	16	11.2	226	1 T21161	Human gene signatu
8	16	11.2	682	1 V69554	U. florida type 1
9	16	11.2	800	1 X39692	Gastric cancer ass
10	16	11.2	1689	1 Q65824	AMEP-1 gene.
11	16	11.2	2132	1 V06392	B. thuringiensis 1
12	16	11.2	2834	1 V61896	Human 14-3-3 zeta
13	16	11.2	3073	1 T14635	Clone associated w
14	16	11.2	3073	1 T34371	Plasmod parv29 (AT
15	16	11.2	8457	1 Q66797	AMEP spheroidin g
16	16	11.2	11000	1 V21209	Methanococcus jann
17	16	11.2	11000	1 X20248	Continuation (9 of
18	15	10.5	132	1 T19753	Human gene signatu
19	15	10.5	192	1 Q85552	DNA probe 58 detec
20	15	10.5	326	1 Q60587	Human brain Expre
21	15	10.5	343	1 T25291	Human gene signatu
22	15	10.5	607	1 V75299	Staphylococcus aur
23	15	10.5	750	1 V53475	DNA encoding a Sta
24	15	10.5	885	1 N20095	Sequence of leukoc
25	15	10.5	888	1 N20010	Hybrid human leuko
26	15	10.5	888	1 Q46120	E. coli 16S rRNA f
27	15	10.5	937	1 N40112	Sequence of the cd
28	15	10.5	937	1 N30105	Sequence on recomb
29	15	10.5	961	1 N40107	Sequence of the cd
30	15	10.5	961	1 N30100	Lap2 promoter seq
31	15	10.5	1177	1 X08427	Schwannomin-blidin
32	15	10.5	1298	1 V32590	16S rRNA gene corr
33	15	10.5	1498	1 Q13005	

C 34	15	10.5	1540	1 N91514	Escherichia coli 1
C 35	15	10.5	1542	1 Q46119	E. coli 16S rRNA f
C 36	15	10.5	1542	1 T18759	E. coli 16S riboso
C 37	15	10.5	1542	1 T29140	rRNA gene (rse) f
C 38	15	10.5	1545	1 Q49060	16S rRNA gene of E
C 39	15	10.5	1743	1 T85939	Shigella flexneri
C 40	15	10.5	1866	1 V53502	DNA encoding a Sta
C 41	15	10.5	2218	1 N90787	Clone L19 of Brass
C 42	15	10.5	2224	1 T40700	N. benthamiana phy
C 43	15	10.5	2273	1 Q04339	Aquaticus I gene
C 44	15	10.5	2274	1 Q12838	Aquaticus I coding
C 45	15	10.5	2274	1 Q75859	Aquaticus I gene.

## ALIGNMENTS

RESULT 1	
T21573	121573 standard; CDNA to mRNA; 84 BP.
AC	T21573;
DT	03-AUG-1996 (first entry)
DE	Human gene signature HUMGS02953.
KW	Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW	human; cloning; mapping; non-biased library; diagnosis; detection;
KW	cell typing; abnormal cell function; ss.
OS	Homo sapiens.
PN	WO9514772-A1.
PD	01-JUN-1995.
PF	11-NOV-1994; J01916.
PR	12-NOV-1993; JP-355504.
PA	(MATS/) MATSUBARA K.
PI	(OKUB/) OKUBO K.
PI	Matsubara K, Okubo K;
DR	WPI: 95-206931/27.
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT	for diagnosis of abnormal cell function, by preparing cDNA that
PT	reflects relative abundance of corresp. mRNA in specific human
PT	tissues
PS	Claim 1: Page 916; 2245pp; Japanese.
CC	A single-stranded DNA (or its complementary strand or the corresp.
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC	given in T19001-T26837 and which is able to hybridise to part of
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC	sequences were obtained from 3'-directed cDNA libraries prepared
CC	from various human tissues; synthesis of cDNA was initiated from the
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC	untranslated sequence is unique to a particular mRNA species, almost
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC	is constructed so as to reflect accurately the relative abundance of
CC	different mRNAs in the particular tissue from which it was derived.
CC	The appearance frequency of a given GS in a cDNA library can be
CC	determined (esp. using primers and probes derived from the GS
CC	sequences) as a means of diagnosing abnormal cell function or for
CC	recognising different cell types.
SQ	Sequence 84 BP; 28 A; 16 C; 19 G; 21 T;
Query Match	14.7%; Score 21; DB 1; Length 84;
Best local Similarity	100.0%; Pred. No. 0.023;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	112 aaaaatgaataaattggtt 132
DB	53 AAAAATTAATTAATTTGTTT 73
RESULT 2	
V74626	121573 standard; DNA; 4121 BP.
AC	V74626;
DT	16-MAR-1999 (first entry)
DE	Staphylococcus aureus coding seq ID #215.

REFERENCE	AUTHORS	JOURNAL	COMMENT
1	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
1	(Bases 1 to 99487)		
1	Sims, S.		
1	Direct Submission		
1	Submitted (18-JUN-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrey@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk		
1	On Feb 16, 2000 this sequence version replaced gi:669798.		
1	IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E. coli, yeast, vector, phage etc. Order of segments is not known; 800's separate segments. Contig-ID: 00017 Length: 1162bp		
1	Contig-ID: 00138 Length: 13765bp		
1	Contig-ID: 00709 Length: 42612bp		
1	Contig-ID: 00828 Length: 10822bp		
1	Contig-ID: 00896 Length: 5150bp		
1	Contig-ID: 00906 Length: 1467bp		
1	Contig-ID: 00963 Length: 27265bp.		
1	* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces		
1	* is not known and their order in this sequence record is		
1	* arbitrary. Gaps between the contigs are represented as		
1	* runs of N, but the exact sizes of the gaps are unknown.		
1	* This record will be updated with the finished sequence		
1	* as soon as it is available and the accession number will		
1	* be preserved.		
1	1162: contig of 1162 bp in length		
1	1163 1962: gap of 800 bp		
1	1963 1572: contig of 13765 bp in length		
1	15728 16527: gap of 800 bp		
1	16528 17811: contig of 1384 bp in length		
1	17912 18724: gap of 800 bp		
1	18712 61323: contig of 42612 bp in length		
1	61324 62123: gap of 800 bp		
1	62124 63205: contig of 1982 bp in length		
1	63206 64005: gap of 800 bp		
1	64006 69155: contig of 5150 bp in length		
1	69156 69955: gap of 800 bp		
1	69956 71422: contig of 1467 bp in length		
1	71423 72222: gap of 800 bp		
1	72223 99487: contig of 27265 bp in length.		
1	Location/Qualifiers		
1	1..99487		
1	/organism="Homo sapiens"		
1	/db_xref="taxon:9606"		
1	/chromosome="6"		
1	/clone_lib="RPCI-3"		
1	/clone_id="RPCI-3"		
1	26800 a 18047 c 17863 g 31206 t 5601 others		
1	BASE COUNT		
1	ORIGIN		
1	Query Match		
1	Best Local Similarity 100.0%; Pred. No. 1.2e+5;		
1	Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0.		
1	OY 126 cccacccaagctcccaaaagctcgtggtatcaagcatgagccac 170		
1	DB 72711 cccacccaagctcccaaaagctcgtggtatcaagcatgagccac 170		
1	AC004977 106474 bp DNA HTG 12-JUN-1998		
1	LOCUS Homo sapiens clone D1152C17, *** SEQUENCING IN PROGRESS ***, 1		
1	DEFINITION unordered pieces.		
1	AC004977		
1	AC004977.1 GI:3213020		

KEYWORDS HTG: HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homidae; Homo.  
AUTHORS 1 (bases 1 to 106474)  
TITLE The sequence of Homo sapiens clone  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 106474)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA

COMMENT \* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

FEATURES  
1 106474: contig of 106474 bp in length.  
1.106474  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DJ1152C17"

BASE COUNT 28957 a 22910 c 23777 g 30830 t  
ORIGIN

Query Match 21.4%; Score 45; DB 41; Length 106474;  
Best Local Similarity 100.0%; Pred. No. 1.2e-15;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 126 cccaccctaagctcccaagtcgtggttcagcgacagccac 170  
|||||  
Db 73539 CCCACCTAAGCTCCCAAGTGTGGATACAGCAGACCCAC 73583

RESULT 12  
HSJ635E18 117859 bp DNA HTG 20-APR-2000  
LOCUS Homo sapiens chromosome 1 clone RP4-635E18 map p36.11-36.31 \*\*\*  
DEFINITION SEQUENCING IN PROGRESS \*\*\* In unordered pieces.  
ACCESSION AL109811.21 GI:7635714  
VERSION AL109811  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 117859)  
AUTHORS Wallis, J.  
TITLE Direct Submission  
JOURNAL Submitted (20-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequests@sanger.ac.uk  
On Apr 22, 2000 this sequence version replaced gi:7242677.  
..... Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
..... Project Information  
Center project name: d635E18  
..... Summary Statistics  
Assembly program: XGAP4; Version 4.5  
Sequencing vector: M13; M77815; 14% of reads  
Sequencing vector: plasmid; 108752; 85% of reads  
Chemistry: Dye-terminator ABI, 1% of reads

Chemistry: Dye-terminator Big Dye; 94% of reads  
Chemistry: Dye-terminator-ambisam; 4% of reads  
Consensus quality: 112388 bases at least Q40  
Consensus quality: 113891 bases at least Q30  
Consensus quality: 114841 bases at least Q20  
Insert size: 116859; sum-of-contigs  
Insert size: 123727; 14.0% error; agarose-fp  
Quality coverage: 5.57x in Q20 bases; sum-of-contigs quality  
coverage: 5.26x in Q20 bases; agarose-fp  
.....

\* NOTE: This is a 'working draft' sequence. It currently \* consists  
\* of 11 contigs. The true order of the pieces is \* not known and  
\* their order in this sequence record is \* arbitrary. Where the  
\* contigs adjacent to the vector can \* be identified, they are  
\* labelled with 'clone.end' in the \* feature table. Some order and  
\* orientation information \* can tentatively be deduced from paired  
\* sequencing reads \* which have been identified to span the gap  
\* between two \* contigs. These are labelled as part of the same \*  
\* 'fragment\_chain', and the order and relative orientation \* of the  
\* pieces within a 'fragment\_chain' is reflected in \* this file. Gaps  
\* between the contigs are represented as \* runs of N, but the exact  
\* sizes of the gaps are unknown. \* This record will be updated with  
\* the finished sequence as \* soon as it is available and the  
\* accession number will be \* preserved.

1 55698 contig of 55698 bp in length  
\* 55799 72183 contig of 16385 bp in length; fragment\_chain 1  
\* 72284 73838 contig of 1555 bp in length; fragment\_chain 1 \*  
\* 73839 75462 contig of 1524 bp in length  
\* 75563 76627 contig of 1365 bp in length  
\* 77028 78261 contig of 1234 bp in length  
\* 78362 79923 contig of 1562 bp in length  
\* 80024 81095 contig of 1072 bp in length  
\* 81196 112651 contig of 1456 bp in length; fragment\_chain 2  
\* 112752 115113 contig of 2362 bp in length; fragment\_chain 2 \*  
\* 115214 117859 contig of 2646 bp in length; fragment\_chain 2 \*  
\* NOTE: This is a 'working draft' sequence  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

FEATURES  
1.117859  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="p36.11-36.31"  
/clone="RP4-635E18"  
/clone\_id="RPCT-4"  
1.55698  
/note="assembly-fragment:00832  
clone\_end:SP6  
vector\_side:left"  
55799..72183  
/note="assembly-fragment:02428  
fragment\_chain:1"  
72284..73838  
/note="assembly-fragment:02464  
fragment\_chain:1"  
73839..75462  
/note="assembly-fragment:00023"  
75563..76927  
/note="assembly-fragment:00489"  
77028..78261  
/note="assembly-fragment:00638"  
78362..79923  
/note="assembly-fragment:00755"  
80024..81095  
/note="assembly-fragment:01032"  
81196..112651  
/note="assembly-fragment:02380  
fragment\_chain:2"  
112752..115113  
/note="assembly-fragment:02462  
fragment\_chain:2"

117: gb-ss813: \*  
118: gb-ss814: \*  
119: gb-ss815: \*  
120: gb-ss816: \*  
121: gb-ss817: \*  
122: gb-ss818: \*  
123: gb-ss819: \*  
124: gb-ss820: \*  
125: gb-ss821: \*  
126: gb-ss822: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	74	16.3	441	22	AA157818	AA157818 2035h07.s
2	71	15.7	417	21	AA099427	AA099427 2179g10.s
3	71	15.7	431	20	AA052890	AA052890 2171e08.s
4	71	15.7	439	21	AA133403	AA133403 2028e05.s
5	71	15.7	447	20	AA055663	AA055663 2175f05.s
6	70	15.5	435	21	AA133390	AA133390 2028b04.s
7	70	15.5	435	21	AA133390	AA133390 2028b04.s
8	69	15.2	501	62	AA007800	AA007800 w03c11.x
9	65	14.3	333	93	AA0045412	AA0045412 RPrC11-32
10	65	14.3	333	93	AA0045412	AA0045412 RPrC11-32
11	65	14.3	333	93	AA0045412	AA0045412 RPrC11-32
12	65	14.3	333	93	AA0045412	AA0045412 RPrC11-32
13	62	13.9	401	41	AA1584112	AA1584112 ts13h03.x
14	61	13.5	439	20	AA034378	AA034378 2K20C10.s
15	61	13.5	439	20	AA034378	AA034378 2K20C10.s
16	59	13.0	289	31	AA807651	AA807651 nx08c07.s
17	59	13.0	292	23	AA243224	AA243224 z12c900.s
18	59	13.0	331	23	AA243224	AA243224 z12c900.s
19	59	13.0	331	23	AA243224	AA243224 z12c900.s
20	59	13.0	331	23	AA243224	AA243224 z12c900.s
21	59	13.0	331	23	AA243224	AA243224 z12c900.s
22	58	12.8	288	31	AA832265	AA832265 oc91f07.s
23	58	12.8	288	31	AA832265	AA832265 oc91f07.s
24	58	12.8	288	31	AA832265	AA832265 oc91f07.s
25	58	12.8	288	31	AA832265	AA832265 oc91f07.s
26	57	12.6	413	90	AA1242047	AA1242047 qh81c06.x
27	57	12.6	413	90	AA1242047	AA1242047 qh81c06.x
28	56	12.4	298	26	AA424412	AA424412 zv83g12.s
29	56	12.4	298	26	AA424412	AA424412 zv83g12.s
30	55	12.1	144	27	AA478483	AA478483 2x16b08.s
31	55	12.1	144	27	AA478483	AA478483 2x16b08.s
32	55	12.1	144	27	AA478483	AA478483 2x16b08.s
33	55	12.1	144	27	AA478483	AA478483 2x16b08.s
34	55	12.1	144	27	AA478483	AA478483 2x16b08.s
35	55	12.1	144	27	AA478483	AA478483 2x16b08.s
36	55	12.1	144	27	AA478483	AA478483 2x16b08.s
37	55	12.1	144	27	AA478483	AA478483 2x16b08.s
38	55	12.1	144	27	AA478483	AA478483 2x16b08.s
39	55	12.1	144	27	AA478483	AA478483 2x16b08.s
40	55	12.1	144	27	AA478483	AA478483 2x16b08.s
41	55	12.1	144	27	AA478483	AA478483 2x16b08.s
42	55	12.1	144	27	AA478483	AA478483 2x16b08.s
43	55	12.1	144	27	AA478483	AA478483 2x16b08.s
44	55	12.1	144	27	AA478483	AA478483 2x16b08.s
45	55	12.1	144	27	AA478483	AA478483 2x16b08.s

## ALIGNMENTS

RESULT 1  
LOCUS AA157818/c  
DEFINITION AA157818 441 bp mRNA MRNA  
2035h07.s1 Stratagene colon (#937204) Homo sapiens CDNA clone  
IMAGE:588925 3' similar to contains LTR7. t3 LTR7 repetitive element

ACCESSION AA157818  
VERSION AA157818.1 GI:1732647  
KEYWORDS EST  
SOURCE human  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 441)  
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaudo, M.F., Chappel, B., Chissey, S., Dietrich, N., Dubucq, T., Favell, A., Gish, N., Hawkins, M., Hulman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rife, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.  
TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
MEDLINE 97044478  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 297.  
FEATURES  
source location/Qualifiers  
1..441  
/organism="Homo sapiens"  
/db\_xref="GDB:462665"  
/db\_xref="taxon:9606"  
/clone="IMAGE:588925"  
/clone\_lib="Stratagene colon (#937204)"  
/issue\_type="tumor"  
/cell\_line="T84 carcinoma cell line"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/note="Organ: colon; Vector: pBluescript SK-; Site: 1; EcorI; Site: 2; XhoI; Cloned unidirectionally. Primer: oligo dt T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'";  
BASE COUNT 116 a 75 c 129 g 119 t 2 others  
ORIGIN  
Query Match 16.3%; Score 74; DB 22; Length 441;  
Best Local Similarity 100.0%; Pred. No. 1.7e-29;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 74 CCTTTCGACTCTCTTTCGACTCAGCCGCTGCACCGATGAATTAACGCT 15  
OY 361 tttgtgtcacaaa 374  
DB 14 TGTGTCTCAAAA 1  
RESULT 2  
LOCUS AA099427/c  
DEFINITION AA099427 417 bp mRNA MRNA  
2179g10.s1 Stratagene colon (#937204) Homo sapiens CDNA clone  
IMAGE:510882 2' similar to contains LTR7. t3 LTR7 repetitive element  
ACCESSION AA099427  
VERSION AA099427.1 GI:1646199  
KEYWORDS EST  
SOURCE human  
ORGANISM Homo sapiens

COMMENT On Jan 6, 2000 this sequence version replaced g1:6676376.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

Seq primer: -400P from Gibco.

# FEATURES

source 1.339

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2309872"

/lab\_host="Soares-NFL-T-GBC-S1"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP-GCB1) were mixed and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

# BASE COUNT

ORIGIN 106 a 87 c 87 g 119 t

# Query Match

Best Local Similarity 100.0%; Score 82; DB 72; Length 399; Pred. No. 1.5e-28; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Qy 1 tgcctcagaagaagcagccctactgacacctgtttggtcgtgtagaccactt 60  
 Db 394 TGCTCCGAGAAAGACGACGCTCTGACACCTGTGGCTGTGGAGACCACTT 335  
 Qy 61 tggactttcactccaact 82  
 Db 334 TGGACTTTTCACCTCCAAACT 313

# RESULT 3

AI076014/c

LOCUS 407 bp mRNA

DEFINITION 27-AUG-1998

3, mRNA sequence.

ACCESSION AI076014

VERSION AI076014.1

KEYWORDS GI:3405192

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 407)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Apr 7, 1998 this sequence version replaced g1:3036064.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

www.bio.lnl.gov/db/tp/image/image.html

Insert Length: 480 Std Error: 0.00

Seq primer: -40m13 fwd. Et from Amersham.

# FEATURES

source 1.407

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:1640467"

/lab\_host="Soares-NFL-T-GBC-S1"

/sex="male"

/lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA

was prepared from mRNA obtained from Clontech

laboratories, Inc., and primed with a Not I - oligo(dT)

primer [5].

TCCTACCAATCTGAGTGGAGGCGGCCCAATTTTTTTTTTTT 3'.

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization to Cot5, and was

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 91 c 88 g 123 t

# Query Match

Best Local Similarity 100.0%; Score 82; DB 35; Length 407; Pred. No. 1.5e-28; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Qy 1 tgcctcagaagaagcagccctactgacacctgtttggtcgtgtagaccactt 60  
 Db 403 TGCTCCGAGAAAGACGACGCTCTGACACCTGTGGCTGTGGAGACCACTT 344  
 Qy 61 tggactttcactccaact 82  
 Db 343 TGGACTTTTCACCTCCAAACT 322

# RESULT 4

AI914390/c

LOCUS 431 bp mRNA

DEFINITION 17-DEC-1999

IMAGE:231365 3', mRNA sequence.

ACCESSION AI914390

VERSION AI914390.1

KEYWORDS GI:5634245

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 431)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

www.bio.lnl.gov/db/tp/image/image.html

Insert Length: 626 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 430.

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:231365"

/lab\_host="Soares-NFL-T-GBC-S1"

/sex="male"

/lab\_host="DH10B"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP-GCB1) were mixed, and ss circles were made in

117: gb-98s13: \*  
118: gb-98s14: \*  
119: gb-98s15: \*  
120: gb-98s16: \*  
121: gb-98s17: \*  
122: gb-98s18: \*  
123: gb-98s19: \*  
124: em-98s13: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	4.4	213	85	H33155 EST108874 R
2	21	4.4	231	33	AA963360 UI-R-EI-g
3	21	4.4	286	33	AA955631 UI-R-EI-f
4	21	4.4	412	33	AA968442 UI-R-CO-1
5	20	4.2	163	93	AA055186 CIT-HSP-2
6	20	4.2	526	37	AA283995 q66e09.x
7	19	4.0	419	47	AL042275 DKFZP434K
8	19	4.0	474	37	AI287912 gnl3c03.x
9	19	4.0	506	47	AI046807 DKFZP434O
10	19	4.0	710	405	AA0482366 RPT-11-2
11	18	3.8	239	91	W85317 mF53H12.11
12	18	3.8	251	63	AA060511 UI-M-BH1-
13	18	3.8	309	93	AA0000680 CIT-HSP-2
14	18	3.8	307	30	AA702247 Z184908.s
15	18	3.8	330	119	AA061388 RPT-23-4
16	18	3.8	352	93	AA0008470 RPT-11-4
17	18	3.8	361	95	AA0200054 RPT-11-4
18	18	3.8	367	40	AA167535 v63c06.x
19	18	3.8	376	102	AA0418263 RPT-11-1
20	18	3.8	387	94	AA0944668 HS-3023B
21	18	3.8	395	29	AA645356 v87912.r
22	18	3.8	420	35	AI073153 UI-R-40-m
23	18	3.8	427	110	AA0752975 HS-5226-B
24	18	3.8	429	120	B38668 HS-7048-A1-
25	18	3.8	446	105	AA0536009 RPT-11-4
26	18	3.8	449	79	AA670115 Z13892 MA
27	18	3.8	458	102	AA0450033 500007C01
28	18	3.8	459	29	AA667436 v17507.r
29	18	3.8	478	95	AA0239927 CIT-HSP-2
30	18	3.8	553	36	AA1180247 EST123990
31	18	3.8	557	79	AA0631901 91282 MAR
32	18	3.8	594	95	AA0234084 HS-2057-A
33	18	3.8	603	47	AA7274271 A7274271
34	18	3.8	614	119	AA0445562 GM-UMD001
35	18	3.8	619	122	Z89665 F. rufipes
36	18	3.8	695	95	AA0194820 RPT-11-65
37	18	3.8	700	96	AA0309735 CIT-HSP-2
38	18	3.8	711	101	AA0330194 nbx0046N
39	18	3.8	715	95	AA0240677 CIT-HSP-2
40	18	3.8	728	47	AA7274313 A7274313
41	18	3.8	740	102	AA0449564 500001H11
42	18	3.8	779	35	AA1096883 qb5f0L.x
43	18	3.8	809	90	W22729 75c5 Human
44	17	3.6	174	92	W219745 HSNAAALEW H
45	17	3.6	183	59	AA042584 AV242584

ALIGNMENTS

RESULT 1 H33155 213 bp mRNA EST 02-APR-1998  
LOCUS H33155 EST108874 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA  
DEFINITION clone RBNAI21 5' end, mRNA sequence.

ACCESSION H33155  
VERSION H33155.1 GI:978572  
KEYWORDS EST  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 213)  
Lee N.H., Weinstein K.G., Kirkness E.F., Earle-Hughes J.A., Puldner R.A., Marmaras S., Glodex A., Gocayne J.D., Adams M.D., Kerlavage A.R., Fraser C.M. and Venter J.C.  
Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment  
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)

JOURNAL MEDLINE  
COMMENT On Jan 6, 2000 this sequence version replaced gi:667654.  
Other ESTs: EST108873  
Contact: Lee, NH  
ATCC  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3829  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
For clone availability please contact the TIGR Database (tigrdb@tigr.org)  
Seq primer: M13 Reverse.

FEATURES  
source  
1..213  
/organism="Rattus sp."  
/db\_xref="ATCC (inhost):2003416"  
/db\_xref="taxon:10118"  
/clone="RBNAI21"  
/clone\_jid="Rat PC-12 cells, NGF-treated (9 days)."  
/note="Vector: pBluescript SK<sup>+</sup>, Site 1: EcoRI, Site 2: XhoI; Poly(A)<sup>+</sup> RNA was purified from 9-day NGF treated PC12 cells. cDNA was constructed using an oligo-dT primer and directionally cloned using the Lambda Zap II Vector kit by Stratagene"

BASE COUNT 73 a 43 c 51 g 44 t 2 others

ORIGIN

Query Match 4.4%; Score 21; DB 85; Length 213;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 gaatgaagaagaagtgaaga 320  
|||||  
Db 2 GAGTGAAGAAGAGGTGAGA 22

RESULT 2  
LOCUS AA963360 231 bp mRNA EST 09-MAR-1998  
DEFINITION UI-R-EI-ga-g-05-0-UI-s1 UI-R-EI Rattus norvegicus cDNA clone  
ACCESSION AA963360  
VERSION AA963360.1 GI:4278284  
KEYWORDS EST  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 231)  
Bonaldo M.F., Lennon G. and Soares M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery  
Genome Res. 6 (9), 791-806 (1996)